Annex C
Modelling of HE achievement

Introduction

1. Annex C has six sections:

- description of the modelling framework we are using to examine HE performance for students (pages 1 to 3)
- description of the principal statistical model we used (pages 3 to 10)
- explanation of the process of calibration for the school performance and school type effects seen in the data. The calibration aims to make the results easier to interpret for the non-statistical reader (pages 10 to 12)
- explanation of the different tests carried out on our modelling procedure in order to confirm (or disprove) the conclusions drawn from it (pages 12 to 21)
- description of the role that multi-level modelling played in our examination of the data (pages 21 to 34)
- references for this annex (page 34).

2. Tables and figures referenced as R(x) appear only in the associated Excel spreadsheet provided with this Word document. Tables and figures referenced as C(x) appear in both the Word document part of the annex and in the associated spreadsheet.

Introduction to the models

<u>Outcomes</u>

3. For each student there is a binary outcome: for example, the student gains an upper second or better by 2001-02; or they do not. Additionally there are a number of potential explanatory variables that could help explain the differences in HE achievement of individuals. We have used logistic regression to model these outcomes. Logistic regression is often used to investigate data where there is a binary response and a set of explanatory variables. Further discussions of logistic regression can be found in many statistical texts, including Collett (1991), and Cox and Snell (1989).

<u>Transfers</u>

4. In a small proportion of cases (around 2 per cent), students change HEIs prior to qualifying during the five-year period. In this study it is assumed that students gain their qualification (if any) from the HEI they began their studies at. This is obviously an over-simplification of the higher education (HE) system, but because we are looking at awards in

a small time period, migration effects are marginal. This means that the model treats a student who began at HEI A and qualified at HEI B, the same as a student who began at HEI A but qualified at HEI A.

Models

5. These data have a hierarchical (or multi-level) structure with, for example, students attending HEIs. Ideally, we would want our models to reflect this structure.

6. As described above, it is assumed that students gain their qualifications (if any) from the HEI they began their studies at, but we know that it is possible for students to move to an alternative HEI during their studies. Multi-level models where lower-level units (students in this case) can belong to two or more higher-level units (for example, attending more than one HEI) are referred to as multiple membership models. Further discussions of these types of models can be found in Goldstein (1995). We do not attempt to fit multiple membership models in this paper.

7. However students also move into HE from a number of post-16 schools. Therefore another multi-level structure separate to the HEI structure could be considered, with students nested within schools. So there are two multi-level structures within the data, with students nested within school, students nested within HEI but, importantly, school not nested within HEI and HEI not nested within school. This is a cross-classification structure.

8. Taking the multi-level structures into account allows us to obtain more statistically accurate estimates of the effects of the explanatory variables, compared against the single-level logistic regression estimates. The information that certain individuals have an association with other individuals (for example, students within the same HEI) also allows for more precise standard errors, and thus a basis for deciding on the statistical significance of explanatory variables. Goldstein (1995) discusses these and other multi-level structures in further detail.

Our approach

9. Given the large number of students involved in our analysis – along with the many HEIs and post-16 schools that students are nested within – fitting the more complex multi-level models without considering the single-level logistic model would be difficult. Testing explanatory variable effects would be time-consuming and may lead to misleading results. So, to correctly model student HE achievement we first modelled the data using logistic regression, extensively examining explanatory variable effects and their relationships with other explanatory variables. The simple logistic model is the main driver of the results reported in the paper. This is a single-level model and can be fitted using any commonly available statistical package; we used SAS because of its effectiveness with large datasets.

10. After modelling the data using logistic regression, we took the multi-level structure of the data into account, with students assumed to be nested within HEIs and also nested within post-16 schools. This type of model is commonly described as a cross-classification

model because there are multiple hierarchical structures within the data that are not nested together. Further multi-level models with a focus on a random coefficient approach are also examined. MLwiN software was used to fit all the multi-level models because of its ability to deal with large and complex multi-level models that other software packages are unable to replicate.

11. Our single-level model was used to form the main results and conclusions of this paper. The key parameters from this single-level model are compared with the parameters from the multi-level models examined, and potential differences caused due to the inclusion of the correct multi-level structures are reported (if any). When there was a chance that a difference conclusion could be formed due to varying results, analyses carried out on the single-level model were repeated for the multi-level model.

12. In general our HE achievement measure was ‘whether the student gained an upper second or higher’. In some analyses, we report the results based on some (or all) of our HE achievement measures (as given in Annex A). However, the main conclusions of the report are derived from the upper second or higher measure.

13. The proportion of students who achieve each outcome, separated by gender, type of school and school performance are given in Annex B, Table B4.

<u>Further models and alternative outcomes</u>

14. Other models using different outcomes could have been fitted to these data, including ordinal regression or a categorical regression approach. Ordinal logistic regression could have been used with degree classification as the single outcome variable. However, interpretation of the results of this type of modelling are difficult to comprehend to a non-statistical reader, and binary outcome models are generally favoured for ease of illustration. Each potential approach/outcome has both advantages and disadvantages over the logistic regression method that we used.

The logistic regression model

<u>The models</u>

Proportion of cohort gaining an upper second or higher by 2001-02

15. Let y_{ijk} be whether student k at HEI i who attended post-16 school j, gains an upper second or higher (1 = yes, 0 = no).

16. There are a number of explanatory variables available to us, of which some are continuous variables and some are categorical. For the logistic model, when the variable is categorical it is separated into dummy variables (that is, each dummy identifying whether the student falls into a particular category). For each categorical variable, all but one of the dummy variables are included. The excluded dummy variable forms the baseline group.

17. Information on the potential explanatory variables is given in Table C1.

Table C1 Explanatory variables for models

Variable type	Explanatory variable	Level	Range / values	Model identifier
Continuous	Student entry qualifications	Individual	(8.0,30.0)	Student Q
	Post-16 school average a-level	School	(0.6,9.0)	School Q
	HEI average a-level pts	HEI	(11.1,29.7)	HEI Q
Category	Gender	Individual	Male	Male
			Female	Baseline
	Degree subject area	Individual	Allied to medicine	Sub 1
			Biological/physical sciences	Sub 2
			Agriculture	Sub 3
			Mathematical sciences	Sub 4
			Engineering	Sub 5
			Social studies	Sub 6
			Business	Sub 7
			Languages	Sub 8
			Creative arts	Sub 9
			Education	Sub 10
			Combined studies	Baseline
	All-girls post-16 school attended	School	All girls	All girls
			Not all girls	Baseline
	Degree course length	Individual	3 year degree	Three year
			4 year degree	Baseline
	School type	School	State	LEA
			Further education	FEC
			Grant maintained	GMS
			Independent	Baseline
	School is selective	School	Selective	Selective
			Not selective	Baseline

18. The structure is then given in the following equation (Equation 1):

$$
\left.\begin{array}{l}
\text{firstup}_i \sim \text{Binomial}(\text{denom}_i, \pi_i) \\
\text{firstup}_i = \pi_i + e_{0i}\text{bcons}_i^*
\end{array}\right\}
$$

$$
\begin{aligned}
\text{logit}(\pi_i) = {} & \beta_1\text{cons}_i + \beta_2\text{student q}_i + \beta_3\text{school q}_i + \beta_4\text{hei q}_i + \beta_5\text{male}_i + \beta_6\text{sub 1}_i + \\
& \beta_7\text{sub 2}_i + \beta_8\text{sub 3}_i + \beta_9\text{sub 4}_i + \beta_{10}\text{sub 5}_i + \beta_{11}\text{sub 6}_i + \beta_{12}\text{sub 7}_i + \\
& \beta_{13}\text{sub 8}_i + \beta_{14}\text{sub 9}_i + \beta_{15}\text{sub 10}_i + \beta_{16}\text{all girls}_i + \beta_{17}\text{three year}_i + \\
& \beta_{18}\text{lea}_i + \beta_{19}\text{fec}_i + \beta_{20}\text{gms}_i + \beta_{21}\text{selective}_i + \beta_{22}\text{student q . school q}_i + \\
& \beta_{23}\text{student q . male}_i + \beta_{24}\text{student q . school q . male}_i + \\
& \beta_{25}\text{student q . sub 1}_i + \beta_{26}\text{student q . sub 4}_i + \beta_{27}\text{student q . sub 5}_i + \\
& \beta_{28}\text{student q . sub 7}_i + \beta_{29}\text{student q . sub 8}_i + \beta_{30}\text{student q . male . sub 5}_i + \\
& \beta_{31}\text{student q . gms}_i + \beta_{32}\text{school q . male}_i + \beta_{33}\text{school q . sub 1}_i + \\
& \beta_{34}\text{male . sub 2}_i + \beta_{35}\text{male . sub 4}_i + \beta_{36}\text{male . sub 5}_i + \beta_{37}\text{male . sub 8}_i + \\
& \beta_{38}\text{male . sub 9}_i + \beta_{39}\text{all girls . lea}_i + \beta_{40}\text{sub 2 . three year}_i + \\
& \beta_{41}\text{sub 8 . three year}_i + \beta_{42}\text{sub 9 . three year}_i + \beta_{43}\text{hei q . three year}_i
\end{aligned}
$$

$$
\text{bcons}_i^* = \text{bcons}_i[\pi_i(1 - \pi_i)/\text{denom}_i]^{0.5}
$$

$$
\left[e_{0i}\right] \sim (0, \Omega_e) : \Omega_e = \left[1\right]
$$

(Equation 1)

where the subscript i refers to student i in the whole dataset (i can vary between 1 and 79,005). The variables bcons and denom are both constant vectors of 1 with size 79,005. The model is specified in this way to allow the data to be modelled within the multi-level software package MLwiN. Interaction terms are denoted with a '.' between each variable within the interaction, so 'student q . school q_i' is the interaction between student qualifications and school quality for student i. Firstup_i is a binary outcome where Firstup_i is 1 when student i has achieved a first or upper second, and is 0 when student i has not.

19. There are 20 explanatory variables that affect the probability of gaining an upper second or better in this equation, and their size is given by β_2 through to β_{21}. There are also 22 interactions between these effects, and their sizes are given by β_{22} through to β_{43}. The remaining β_1 sets the constant effect in this logistic equation.

20. Alternative models that include additional (or substitute) effects, including interactions and quadratic terms, have been examined. The exclusion of some currently included effects cause other currently excluded effects to become significant, and extensive exploratory work has been carried out to confirm that these alternative models do not cause our conclusions to change.

21. For example, there appears to be a quadratic nature to the student qualification effect, but this seems to be partly down to the relationship between school quality and student qualifications, and also due to a student's A-level points being truncated at 30. The

conclusions drawn from the model structure in Equation 1 still hold for the alternative models examined.

22. Table C2 shows the parameter, significance and meaning of each β term in the logistic equation.

Table C2 Parameter estimates for original model

Beta	Effect of	Parameter	P-value
β1	The constant	-1.636	0.000
β2	Student A-level pts	0.107	0.000
β3	School average A-level pts	-0.114	0.001
β4	HEI Average A-level pts	-0.012	0.003
β5	Being male	0.614	0.017
β6	Studying subjects allied to medicine	0.953	0.000
β7	Studying biological or physical sciences	-0.120	0.029
β8	Studying agriculture	0.464	0.000
β9	Studying mathematical sciences	0.420	0.000
β10	Studying engineering	1.249	0.000
β11	Studying social studies	-0.123	0.000
β12	Studying business	0.314	0.000
β13	Studying languages	-0.753	0.000
β14	Studying creative arts	-0.204	0.151
β15	Studying education	0.039	0.374
β16	Attending an all-girls school	0.152	0.000
β17	Being on a three-year course	-0.809	0.000
β18	Attending a state-school post-16	0.344	0.000
β19	Attending a FEC post-16	0.195	0.000
β20	Attending a grant-maintained school post-16	0.382	0.000
β21	Attending a selective school	-0.085	0.007
β22	(Additional) student A-level pts combined with school A-level pts	0.007	0.000
β23	(Additional) student A-level pts on males	-0.029	0.023
β24	(Additional) student A-level pts crossed by school A-level pts on males	0.007	0.003
β25	(Additional) student A-level pts for subjects allied to medicine	-0.018	0.008
β26	(Additional) student A-level pts for mathematical sciences	-0.039	0.000
β27	(Additional) student A-level pts for engineering	-0.064	0.000
β28	(Additional) student A-level pts for business	-0.010	0.032
β29	(Additional) student A-level pts for languages	0.028	0.000
β30	(Additional) student A-level pts for males studying engineering	0.033	0.014
β31	(Additional) student A-level pts for grant maintained post-16	-0.008	0.026
β32	(Additional) school A-level pts for males	-0.240	0.000
β33	(Additional) school A-level pts for subjects allied to medicine	-0.088	0.019
β34	(Additional) being male and studying biology/physics	-0.198	0.000
β35	(Additional) being male and studying mathematical sciences	0.155	0.022
β36	(Additional) being male and studying engineering	-0.853	0.002
β37	(Additional) being male and studying languages	0.218	0.000
β38	(Additional) being male and studying creative arts	0.248	0.006
β39	(Additional) attending an all-girls state-school post-16	-0.119	0.044
β40	(Additional) being on a three-year course and studying biology/physics	0.245	0.000
β41	(Additional) being on a three-year course and studying languages	0.306	0.000
β42	(Additional) being on a three-year course and studying creative arts	0.363	0.011
β43	(Additional) HEI A-level pts on those on a three-year course	0.028	0.000

23. Table C2 shows that all but two of the variables included in this model are statistically significant. The two non-significant variables are included for completeness due to lower-level interaction terms.

24. A significance level of less than 0.05 indicates that the variable is statistically significant. A p-value of less than 0.01 shows that the variable is highly significant, and less than 0.001 means that the explanatory variable has an extremely high significant effect on the probability of gaining an upper second (or higher) at degree level.

25. Imagine there is a student with the following characteristics:

- male
- achieved 28 points at A-level
- post-16 studied at an selective independent school, whose average grade per A-level entry was a B (eight points)
- studying mathematics at a HEI, whose average entry requirement was 24.8 A-level points and was on a four year course.

26. The model would calculate the probability of this student gaining a first or upper second within five years of entry in the following way:

z (the probability of gaining a first or upper second) = exp(α) / (1 + exp(α)),

where α =	-1.636	*The constant*
	+ 0.107 * 28	Student had 28 A-level points on entry
	- 0.114 * 8	School average points per A-level entry
	- 0.012 * 24.8	HEI has an average entry requirement of 24.8 points
	+ 0.614	Student is male
	+ 0.420	Student is studying mathematical sciences
	- 0.085	Attended a selective school
	+ 0.007 * 28 * 8	Relationship of student and school A-level points
	- 0.029 * 28	Additional effect of student A-levels for a male
	+ 0.007 * 28 * 8	Additional effect of relationship between student and school A-level points for a male
	- 0.039 * 28	Additional effect of student A-levels for a mathematics student
	- 0.240 * 8	Additional effect of school average points per A-level entry for a male
	+ 0.155	Additional effect of being male on a mathematics course
=	0.523	

(Equation 2)

27. This means that the probability of gaining a first or upper second for this student would be equal to exp (0.523) / (1 + exp (0.523)) = 62.8%.

28. The interactions within the model make it difficult to interpret the differing effects of student qualifications and school performance for different types of student. The Excel spreadsheet with this annex uses a contour plot to illustrate the estimated outcomes for

different types of individuals. The spreadsheet allows the user to vary the student type, and shows how a student's A-level points and the performance of the school they attended affects their probability of achieving an upper second or higher. The spreadsheet allows the user to modify the following characteristics:

- the average A-level entry requirements of the HEI the student attends. This can vary between eight A-level points (for example, an HEI whose average student A-level attainment is DD) and 30 A-level points (AAA)
- the subject of the degree the student is taking at the HEI; these fall into 11 general categories
- whether the student is male (M) or female (F)
- whether the student attended an LEA, further education college, grant maintained or independent school for their post-16 education prior to entry to the HEI
- whether the school the student attended prior to HEI entry was a selective school
- whether the school the student attended prior to HEI entry was an all girls school
- whether the expected length of the degree course being taken is three or four years long.

29. The contour map in the spreadsheet shows how the probability of gaining an upper second or higher (z-axis) varies by the student's A-level points (x-axis) and the performance of the school the student attended (y-axis).

Other outcomes with a linear relationship with student A-level points

30. The relationship between student A-level points and the probability of gaining an upper second or above is approximately linear, as seen in Figure 1 of the main report. The same general linear relationship is seen for three out of four of the other student outcomes examined:

- proportion of cohort who gained an HE award by, or are still active in, 2001-02
- proportion of cohort who gained an honours degree by 2001-02
- proportion of cohort who gained a lower second or better by 2001-02.

31. For these three outcomes, the same effects as in the upper second or above model have been examined. This means using exactly the same model structure as given in Equation 1 but for each of the three outcomes in turn. The parameter estimates of the effect on each outcome is allowed to vary, and Table R1 shows the parameter estimates alongside their statistical significance on each outcome.

32. In each model, some terms are statistically significant and some are not. In a large proportion of cases, the direction of the effect of each variable is either the same for each model or the variable is not significant, that is, the effects are consistent for all four models. The effect of student A-level points is always significant and always positive. This means that for all four outcomes, the more points the student has, the higher the probability of a 'successful' outcome.

First class honours only

33. Figure 1 in the main report shows that the relationship between the probability of gaining a first class honours degree and a student's A-level points is non-linear. This is a significantly different pattern compared to the relationship between a student's A-level points and the other four outcomes.

34. The same general structure applies as before but the outcome of interest is now first_i, which is 1 when student i has achieves a first and 0 when they have not. It is important to note that the beta terms do not necessarily apply to the same effects as seen in Equation 1. The results using 'first class honours only' as the student outcome are given in Equation 2.

$$\left.\begin{array}{l}\text{first}_i \sim \text{Binomial}(\text{denom}_i, \pi_i) \\ \text{first}_i = \pi_i + e_{0i}\text{bcons}_i^*\end{array}\right\}$$

$$\begin{aligned}\text{logit}(\pi_i) = {} & \beta_1\text{cons}_i + \beta_2\text{student q}_i + \beta_3\text{school q}_i + \beta_4\text{hei q}_i + \beta_5\text{male}_i + \\ & \beta_6\text{sub 1}_i + \beta_7\text{sub 2}_i + \beta_8\text{sub 3}_i + \beta_9\text{sub 4}_i + \beta_{10}\text{sub 5}_i + \beta_{11}\text{sub 6}_i + \\ & \beta_{12}\text{sub 7}_i + \beta_{13}\text{sub 8}_i + \beta_{14}\text{sub 9}_i + \beta_{15}\text{sub 10}_i + \beta_{16}\text{all girls}_i + \\ & \beta_{17}\text{three year}_i + \beta_{18}\text{lea}_i + \beta_{19}\text{fec}_i + \beta_{20}\text{gms}_i + \\ & \beta_{21}\text{student q . school q}_i + \beta_{22}\text{student q . male}_i + \\ & \beta_{23}\text{student q . school q . male}_i + \beta_{24}\text{student q . sub 1}_i + \\ & \beta_{25}\text{student q . sub 2}_i + \beta_{26}\text{student q . sub 4}_i + \beta_{27}\text{student q . sub 5}_i + \\ & \beta_{28}\text{student q . sub 7}_i + \beta_{29}\text{student q . sub 9}_i + \beta_{30}\text{student q . sub 10}_i + \\ & \beta_{31}\text{school q . male}_i + \beta_{32}\text{male . sub 5}_i + \beta_{33}\text{male . sub 7}_i + \\ & \beta_{34}\text{male . sub 8}_i + \beta_{35}\text{male . sub 9}_i + \beta_{36}\text{three year . hei q}_i + \\ & \beta_{37}\text{sub 1 . three year}_i + \beta_{38}\text{sub 2 . three year}_i + \\ & \beta_{39}\text{student q . student q}_i + \beta_{40}\text{student q . student q . sub 7}_i + \\ & \beta_{41}\text{student q . student q . sub 10}_i\end{aligned}$$

$$\text{bcons}_i^* = \text{bcons}_i[\pi_i(1 - \pi_i)/\text{denom}_i]^{0.5}$$

$$\left[e_{0i}\right] \sim (0,\ \Omega_e) : \Omega_e = \left[1\right]$$

(Equation 3)

35. Some new terms are included for the first-class only model. Additionally some β terms have been excluded in comparison to the models given in Equation 3. This is because they are non-significant and not key in maintaining the interaction validity of the model. Full parameter estimates for each beta are given in Table R2.

36. The inclusion of additional interaction terms and exclusion of others makes it difficult to directly compare the differences between the first-only model and the other models. For example, looking at the raw effect of student qualifications it appears that the direction of the effect is reversed for the first-only model (-0.09) compared to values of around +0.09 for the other models. This is mainly due to the inclusion of a squared term in the first-only model for student qualifications, which is not present in the other four models.

Calibrating effects

Calibrating school performance effects

37. In the remainder of this annex, the logistic regression model that uses upper second or higher as its outcome will be the main focus. This model will be referred to as the original model.

38. Due to the large number of significant interactions relating to school performance, it is difficult to conceptualise its effects on the outcome of interest. To understand school performance more clearly for each individual, we modify a student's school performance by some set amount and then adjust the A-level points of that student so that they have the same probability of a successful outcome based on their original (and actual) characteristics.

39. In modelling terms, the process of calibration for schooling effect differences using A-level points is described below. For each individual in the actual data, calculate the probability of a successful outcome using Equation 1 and the beta parameters given in Table C2. Then reduce each individual's actual school performance by 0.1 of a point. The selectiveness of the school remains the same as in the original data.

40. Using the same beta terms given in Table C2, we could calculate a new probability of success for each modified individual. But rather than do this, we fix the probability of success to what it was originally and vary the student's A-level points. We now know everything apart from one unknown, student A-level points. Simple algebra allows us to rearrange the model equation given in Equation 1 to provide us with an estimate of what the student A-level points would have to be to achieve the same probability of success, given the modified school quality.

41. For each student, we now have the required change in A-level points to achieve the same probability of success as a student who went to a school which was 0.1 points better in terms of our quality measure. The change in points is multiplied by 20 to create a scale based on a modification in school quality of 2 points rather than 0.1 points.

Defining school performance groups

42. In Table 1 of the main report, schools are placed into four different performance groups. The method is as follows:

- Take all the students in the population and sort them in ascending order by their school performance (as measured by the average points achieved per A-level entry for that school).
- Allocate the schools of those students who fall into the first 25 per cent of the sorted data to the lowest performing school group.
- All students who attend the same school will be allocated to the same school performance group. This means that if a student falls outside the first 25 per cent of sorted students but attends the same school as a student who falls within the first 25 per cent, they will be allocated to the lowest school performance group.
- Repeat the process for the students who fall in between the 25 and 50 per cent in the sorted data, and who have not yet been allocated to a school performance range.
- Repeat for the students who fall between 50 and 75 per cent of the sorted data.
- All remaining students have schools that fall into the highest school performance range.

Calibrating school type effects

43. The way in which the school type attended by the individual affects HE achievement is difficult to interpret due to the large number of interactions, similar to the problems with interpreting school performance. Given this, it seems obviously that a similar approach to that of interpreting school performance should be taken. This means, rather than taking 0.1 off a student's type (which wouldn't make statistical or practical sense), we can 'flip' a student's school type. In the original data each individual has a school type but we could convert all school types to, say, state school for all individuals and do similar calculations as in paragraphs 40 and 41, but replacing school performance modifications with school type modifications. This would mean that all individuals who attended state schools in the original data would be unaffected by any changes as their school type is being 'flipped' from state school to state school. The selective nature of the school is unchanged.

44. This idea isn't a particularly close representation of the real world because the characteristics in performances of each school type are so different. It is rare to find a very low performance independent school (say averaging E or less for each A-level entry). Also it is rare to find a state school whose average is well above a B grade. So changing some individuals' school types would create artificial schools that are very unlikely to exist in the real world and thus would give a bias (and incorrect) feel for the effect of school type.

45. To achieve a more realistic examination of the effect of school type on a student outcome, we simultaneous modify the school type and the school performance. The changed school performance for each individual is assigned to be a random selection from

the school performances of the modified school type. For example, if all individuals have their school type modified to state school, then the school performances of each individual would be set to a randomly selected state school performance from the original data.

46. As before, the only unknown for the modified data is the student A-level points score and this can be inferred using the method described in paragraphs 40 and 41. The school performance is not changed for those individuals whose original school type is the same as the modified school type. This process is repeated a number of times, and the mean of the simulations is taken for the parameters of interest. Fifty simulations were carried out.

47. Table 3 in the main document shows the effect of the flip for LEA students who are changed to being independent school students. For this table, the school performance is not modified.

48. Table 4 in the main document shows the same flip (that is, LEA to independent school) but with the school performance of the flipped student changed to a performance that is in line with an independent school's performance.

49. Table C3 shows the effect of 'flipping' a student school type and school performance for all possible combinations of flips, that is, the equivalent to Table 4 but for all possible school type flips for an average student from that school type. The table gives the increase in A-level points required for the student to have to same probability of HE achievement as when they attended their original school type. For example, an average FEC student will require roughly two additional A-level points to reach the same HE achievement if they had attended an independent school instead of a FEC. The value of three A-level points given in the table for an LEA to independent school flip corresponds to a weighted average of the students in Table 4 from the main report.

Table C3 Effect of changing school type and school performance simultaneously

Original school type	Changed school type Mean increase in HE achievement (percentage of students increasing)			
	LEA	FEC	Direct grant	Independent
LEA		1 (7%)	1 (12%)	3 (0%)
FEC	-1 (91%)		0 (62%)	2 (4%)
Direct grant	-1 (75%)	0 (34%)		2 (4%)
Independent	-2 (75%)	-1 (85%)	-1 (86%)	

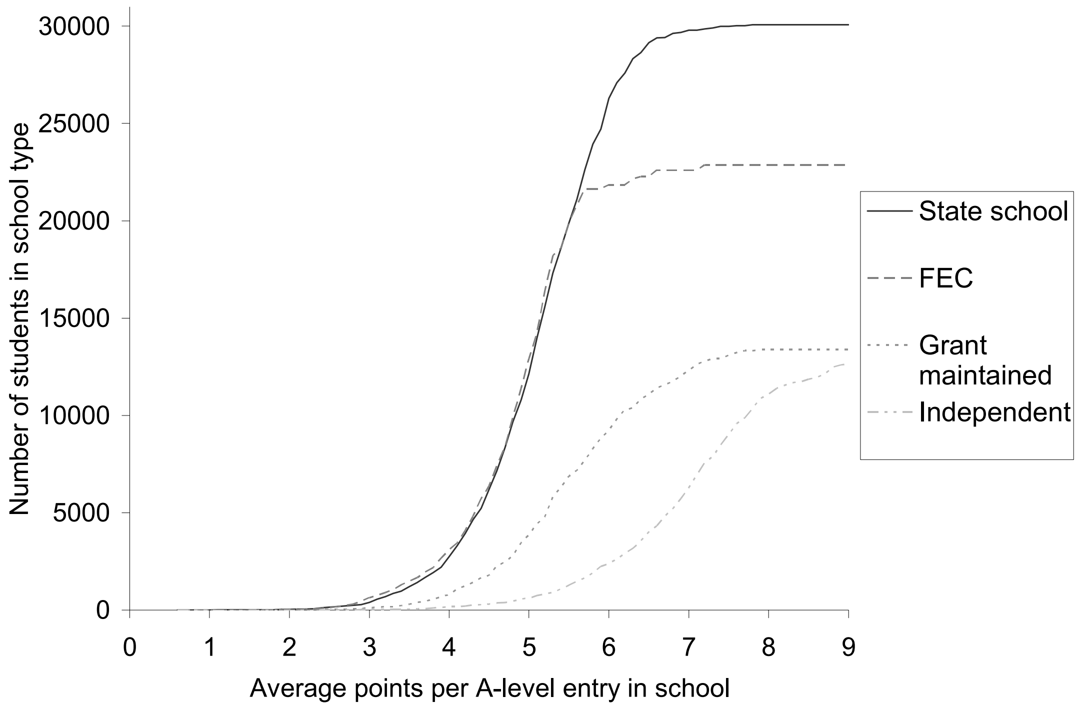
50. Tables 3 and 4 in the main report and Table C3 are based on gaining an upper second or better. The equivalent results to Table 3 from the main report for the other four outcomes (alongside the original Table 3) are given in Table R3.

Testing the model

Introduction

51. Figure C1 shows the cumulative distribution of school performance for each school type (weighted by student numbers). It shows that a significant proportion of students from schools with higher average A-level points per entry attend independent schools. Additionally all schools with average points per A-level entry of 8.0 or over (i.e. averaging a B per A-level entry) are independent. This means that there is a high co-linearity between school performance and school type. It is possible that school performance effects could be masked by school type effects, that is, the school type could be acting as a pseudo variable for school performance.

Figure C1 Cumulative school performance distributions for each school type



52. Three single-level model approaches are used to test whether school-type could be acting as a proxy for school performance. A series of multi-level models have also been fitted to examine this and other issues, which are described later in this annex. The three single-level model tests are:

- extending the descriptive statistics
- elimination of high performing schools from the dataset
- generating simulations where school type is set so it does not affect student outcomes.

Extending the descriptive statistics

53. Figure 2 from the main report – which shows how HE achievement varies for students with different A-level points and who attended schools of differing performance – is reproduced for different sub-sets of students: Figure R1 shows the equivalent plot to Figure 2 for all females in the data; Figure R2 does the same for males only; and Figures R3-R6

show the effect for males who attended the LEA, FEC, grant-maintained and independent school types respectively.

54. The plots show that the differences in HE achievement between students disappear for female students; that is, for female students with the same A-level points, the performance of their school seems to have little effect on their HE achievement. For male students considered as a whole a difference in the school performance types still exists, with the highest performing school students gaining the lowest HE achievement by A-level points. This consistently low HE achievement for the highest performing schools disappears when males of varying school types are examined separately.

Elimination of high performing schools

55. If we eliminate schools above a certain school performance level, then the confounding between school performance and school type is reduced; that is, the probability of those schools from the higher ranges of school performance being independent is reduced and a better spread of school types is produced. The models can be re-run with these reduced data, and if the school type effects are dramatically reduced then this gives some indication that school type could be acting as a proxy for school performance.

56. The analysis was re-run with two different upper limits: 5.0 and 6.0.

Upper limit: 6.0

57. There were 59,752 students whose school performance was 6.0 average A-level points per exam entry or less. They were made up of: 26,289 state school students; 21,851 further education college students; 9,224 grant-maintained school students; and 2,388 independent school students.

58. The same analysis that produced Tables 3 and C3 can be carried out on these truncated data. Table C4 shows the truncated data equivalent to Table 3 and Table C5 to Table C3.

Table C4 Truncated data equivalent (6.0 and below) to Table 3

Individual student A-level points	HE achievement – A-level point equivalents	
	Males	Females
low (5–8)	3.5 (100%)	1.6 (100%)
10	3.2 (100%)	1.2 (100%)
12	2.8 (100%)	0.9 (100%)
14	2.5 (100%)	0.5 (100%)
16	2.1 (100%)	0.2 (100%)
18	1.8 (100%)	-0.2 (6%)
20	1.5 (100%)	-0.5 (8%)
22	1.2 (100%)	-0.8 (8%)
24	0.9 (100%)	-1.2 (8%)

26	0.6 (100%)	-1.5 (0%)
28	0.2 (100%)	-1.9 (0%)
30	-0.1 (1%)	-2.2 (0%)

Table C5 Truncated data equivalent (6.0 and below) to Table C3

Original	Changed school type			
school type	LEA	FEC	Direct grant	Independent
LEA		1 (9%)	1 (15%)	3 (1%)
FEC	-1 (92%)		0 (62%)	2 (2%)
Direct grant	-1 (77%)	0 (36%)		2 (2%)
Independent	-2 (98%)	-2 (96%)	-2 (96%)	

59. The same parameters used in the simple logistic model (student success being achieving an upper second or above) are fitted for these data. The estimates and significance of these parameters for these truncated data are given in Table R4.

Upper limit: 5.0

60. There were 29,564 students whose school performance was 5.0 average A-level points per exam entry or less. They made up of: 12,141 state school students; 12,935 further education college students; 3,873 grant-maintained school students; and 615 independent school students.

61. The equivalent tables to Tables 3 and C3 are given in Table C6.

Table C6 Truncated data equivalent (5.0 and below) to Table 3

Individual student	HE achievement – A-level point equivalents	
A-level points	Males	Females
low (5–8)	3.3 (100%)	2.4 (100%)
10	3.2 (100%)	1.8 (100%)
12	3.0 (100%)	1.2 (100%)
14	2.9 (100%)	0.7 (100%)
16	2.8 (100%)	0.1 (100%)
18	2.7 (100%)	-0.4 (5%)
20	2.6 (100%)	-0.8 (8%)
22	2.5 (100%)	-1.3 (8%)
24	2.3 (100%)	-1.9 (8%)
26	2.2 (100%)	-2.4 (0%)
28	2.2 (100%)	-3.0 (0%)
30	2.0 (100%)	-3.5 (0%)

Table C7 Truncated data equivalent (6.0 and below) to Table C3

Original school type	Changed school type			
	LEA	FEC	Direct grant	Independent
LEA		1 (14%)	0 (33%)	2 (6%)
FEC	-1 (89%)		-1 (75%)	1 (5%)
Direct grant	0 (63%)	1 (24%)		2 (3%)
Independent	-2 (79%)	-1 (93%)	-2 (97%)	

62. The parameter estimates using these truncated data are given in Table R5.

Conclusions from the elimination approach

63. The analysis of the effects of school performance and school type for the truncated data produce tables which should similar patterns to the original data. The same conclusions can be drawn from these new tables as were drawn from the original analysis. There is no reduction in the independent school effect.

64. The relative effect of school performance compared to the school type is stronger in the truncated data. This is mainly because we have purposely modified the data to attempt to eliminate any confounding effect of school type on school performance, that is, the signal for school performance is clearer as school type is not very highly correlated with it anymore.

No school type effects

Using the same effect as seen in the original model

65. A major concern of the recorded school type effects is that they are really school performance effects being picked up by the high correlation between school type and school performance. In the previous section we examined the results of truncating the original data to reduce this effect, but another approach would be to create a scenario where there are known to be no school type effects but significant school performance effects.

66. Each student in the original dataset has a predicted probability of success based on the parameters in Equation 1. We can modify this equation so that the predicted probability of success does not depend on school type but still depends on all the other factors in the model. This means removing those terms that involve school type from the original equations, which means that the following β terms are removed: β_{17}, β_{18}, β_{19}, β_{29} and β_{37}. Assuming that the other effects have the same size as recorded in the original probability equation, we can calculate a new probability of success which, by design, does not depend on school type.

67. Using this approach, we can generate an outcome for each student based on this new probability of success that is not dependent on school type. This is then used in place of the original outcome, and our original single level modelling and analyses are re-run. Given that the outcomes for each student are generated randomly, another set of randomly generated outcomes may produce a different set of results. This means that we need to repeat the process of generating outcomes for each student and analysing the results of each of the

generated datasets. The school type of each individual is also regenerated for each simulation run.

68. Table R6 shows the mean parameter estimates for each effect tested for using around 50 simulated data runs and their associated Monte Carlo standard errors. Additionally the number of times that effect was determined to be significant in the simulations is given. It shows that the pre-determined target effects are identified with reasonable accuracy using the original modelling. Those school type variables that were pre-determined to be zero are only found to be significant around 5 per cent of the time using the original modelling techniques (as expected due to random variation). In around 11 per cent of the simulations, the school average A-level points were found not to have a significant effect. The concern would be that for these occurrences, the school type variables take the place of the school point effect, but this is not the case. In only one of these cases does a main school type variable become significant, and the recorded effect for this school type variable is much smaller than is seen in the actual results.

Modifying the pre-determined effect of school performance

69. As we have pre-determined the effect of school performance and all the other student characteristics, it is possible to modify them to create stronger (or weaker) effects of these parameters on student outcome. The effects of modifying school performance by making it stronger and weaker were investigated. The same conclusions as with using the same effects as seen in the original model were reached in these analyses: the school type effects given in the original model could not be generated using pre-determined school performance effects.

Model effects for different clusters of HEIs

70. One concern of the model is that although it can correctly identify effects at a sector-wide level, it does not adequately describe effects at a lower level. Effects noted for the general population may not apply at an HEI level for example. To examine whether there is significant differences for certain parts of the HEI population, we separate the HEIs into five different clusters. Each cluster contained HEIs with similar entry requirements, based on A-level points of the students who enter. A cluster analysis technique was used to determine these clusters. Table C8 shows the number of HEIs in each cluster and the average HEI entry requirements. Table R7 indicates which cluster each HEI falls into.

Table C8 Information on HEI performance clusters

Institution groups	Number of institutions	Average HEI A-level points
1 (Highest)	4	28.6
2 (2nd highest)	12	25.1
3 (3rd highest)	23	21.6
4 (4th highest)	34	16.7
5 (Lowest)	44	13.6

71. Tables 1 and 2 are reproduced for each cluster and are given in Tables R8-R12 (lowest performing HEI cluster in Table R8, through to highest performing cluster in Table R12). For the reproduction of Table 1, the school performance ranges are now based solely on the schools within that cluster of HEIs.

72. In the lowest performing cluster, the model underestimates the proportion of firsts and upper seconds by around 0.5 per cent. It overestimates this proportion for the fourth and third highest clusters, and for the highest cluster it underestimates again. This gives some evidence that there is a quadratic relationship between HEI quality and the proportion of upper seconds or higher, as the model is underestimating at the edges of the range of HEI qualities and overestimating in the middle. This pattern and others noted in the residuals indicate that additional terms should be considered in the model.

73. Upon inspection, this quadratic behaviour seems to be partly down to the performance of students recorded as having 30 A-level points. These students are performing particularly well in comparison to the rest of the population, and the highest performing clusters have the large majority of these students. The effect of these students also varies by HEI, and this implies that two additional terms should be included in any updated model: a dummy variable for students with 30 A-level points; and an interaction between this new term and HEI quality. This updated model is given in Table R13 and the results are described below.

74. The main conclusions drawn from this updated model are the same as the original modelling.

75. The performance distribution of schools for the highest performing cluster is significantly higher than for the other clusters. The median school performance for a student in the highest performing cluster is 6.5, that is, between a C and B grade per A-level entry. The next cluster has a median of 5.6, which equates to below a C grade per A-level entry from each school. The highest cluster also takes a large proportion of independent school pupils in comparison to the other clusters. The highest performing cluster is the only one where independent school pupils are in the majority.

76. In all of the clusters, the independent school pupils do worse than the LEA students in terms of proportion achieving an upper second or higher, but their average A-level points are always higher.

77. There is some variation in the behaviour of students between the clusters. Comparing the actual results of students to their expected results can help identify these variations. For example, the expected difference between the lowest performing school quartile and the highest performing school quartile is predicted to be just below 4 per cent for the highest HEI cluster, but in the actual data the difference is only 2 per cent.

78. The tables show that the majority of the overall conclusions hold for all of the HEI clusters but some differences can be seen in certain clusters.

Updated model based on residual analysis

79. The result of the analysis of the residuals for the different clusters of HEI shows that additional terms may be appropriate in the modelling. The inclusion of these additional effects increases the model complexity but also significantly improves the model's explanation of the variation in the data. The size and effect of both the original and new terms for this updated model are given in Table R13.

80. Tables 3 and C3 are re-created using the updated model and are given in Tables R14 and R15. The original school performance effects noted in Table 3 for females have now been dramatically reduced, and can be explained by alternative effects included in the modelling. The school performance effects for males are also reduced to a lesser extent.

81. In the original modelling, school performance effects lacked consistency. In the updated model, school performance effects become even less consistent and are reduced. The comparison of Table C3 to its equivalent for the updated model shows that the school type effects are unchanged. The same patterns are noted with independent school pupils performance varying by around three points when compared to similar state school pupils.

82. The residual analysis based on clusters of HEIs (the same clusters as were assumed in the original analysis) shows that the trend for under-prediction of degree performance for the higher clusters has disappeared. The quality of HEIs within clusters is now not significantly correlated with the model residuals.

83. There are many other alternatives for modelling the effect noted for students with 30 A-level points, and some of these have been explored. The results are not reported in this document, but the main conclusions produced from the original modelling relating to school performance and school type remain unaffected for the models examined.

84. One alternative approach would be to allow β_{42}, the parameter for the dummy variable for students with 30 A-level points, to vary at an HEI level. This would mean that this beta constant would have a fixed effect part showing the overall additional effect of having 30 A-level points on the outcome. It would also allow for a different random effect term for each HEI, meaning that a 30 point student at HEI A can have a different effect to a similar 30 point student at HEI B. Distributional assumptions regarding the HEI variation for this beta variable must also be made. For example, the HEI variation could be assumed to come from a normal distribution, even that the HEI variation must be positive. In general, such models would exclude the interaction between the HEI quality and the β_{42} term that was included in the updated model described previously.

85. Other alternative approaches could include fitting latent terms at a student level which would be included in the student's A-level qualifications if the student had 30 A-level points. Each 30 point student's entry qualifications would be assumed to have the form ($30 + \mu_{ijk}$), with distributional assumptions placed upon the μ_{ijk}'s (for example, they come from a non-negative distribution). Such models would help to examine in more detail how the variation is partitioned between the levels in the data, the covariates and misclassification in the

covariates (for example, the effect of truncating student A-level points at 30 for the data we are using).

Weighted analysis

86. One feature of our data is that as an HEI's average A-level points per entry rises, the proportion of students who attend from an independent school also rises. This association could be partly driving the school type effects seen in the original modelling. Independent school students tend to go to more selective HEIs, and at more selective HEIs it is harder to gain a good degree. This would mean that independent school students would appear to do worse than their alternative school type counterparts because they gain their degrees from HEIs where it is harder to gain a good degree.

87. Our HEI quality variable included in the original modelling should help adjust for this effect. It could be argued that this HEI quality variable could be highly confounded with the proportion of independent school entrants at the HEI.

88. The cluster analysis, which examined HEIs with similar levels of quality, should also take this confounding into account. The same trend in school type effects are seen within HEIs of the same general quality, that is, the HEI quality effects have been negated as the cluster analyses only examines HEIs of similar quality.

89. One other potential way of adjusting for the confounding between HEI quality and the proportion of independent school students attending is to perform a weighted logistic regression of the data.

90. In the overall population, there is an underlying proportion of independent school students. Our target dataset is one in which this proportion remains constant across all HEIs, in other words, the proportion of independent school pupils is the same at each HEI. To achieve this we will weight each observation in our dataset so that observations for state school students will have more weight at HEIs where there is a large proportion of independent school students. A logistic regression can then be performing using the weighted observations.

91. The weights are allocated as follows:

- Let $p_.$ be the proportion of independent school students in the whole dataset.
- Let p_j be the proportion of independent school students within HEI j.
- Let w_{1j} be the weight allocated to an independent school student at HEI j.
- Let w_{2j} be the weight allocated to a non-independent school student at HEI j.
- If p_j is not equal to 0 or 1 then $W_{1j} = p_+ / p_j$, and $W_{2j} = (1 – p_+) / (1 – p_j)$, else $W_{1j} = W_{2j} = 1$.

92. Under these conditions, each HEI would still contribute the same number of students as they did in the original analysis, but the effects of different proportions of independent school pupils at different HEIs should be negated. In the case when the cell has all independent (or conversely all non-independent) school students, the weights are set to 1 as

both types of student (non-independent and independent) are required to correctly weight a cell.

93. The parameter estimates and their significances from the original model, but now using the weighted observations, are given in Table R16. The same weighted analysis can be re-calculated by setting the proportions within subject areas at each HEI to be set to the overall population proportions. The calculations are as before but the base cell is now assumed to be 'HEI by student's qualification' rather than 'HEI'. The structure is now

- Let p_{jk} be the proportion of independent school students within HEI j with student qualification k (for example 29-30 A-level points)
- Let w_{1jk} be the weight allocated to an independent school student at HEI j with student qualification k
- Let w_{2jk} be the weight allocated to a non-independent school student at HEI j with student qualification k
- If p_{jk} not equal to 0 or 1 then $W_{1j} = p_+ / p_{jk}$, and $W_{2j} = (1 – p_+) / (1 – p_{jk})$, else $W_{1j\,k} = W_{2j\,k} = 1$

94. The results of this 'HEI by student's qualification' are given in Table R16. The results of three other additional analyses are also reported in this table, where the proportions are fixed within: HEI and subject area; HEI and gender; and HEI by gender, subject area and student's qualification.

95. Table R16 shows that for all of the weighted analyses the relationship between state school and independent school pupils remains consistent. The beta parameter (β_{18}) remains at around 0.35. This finding is consistent with the theory that the school type effects noted in the original unweighted model are not due to the confounding between proportion of independent school student at a HEI and the entry A-level requirement of the HEI.

Multi-level models

Cross-classification model

96. As discussed in paragraphs 5-8, it is important not to ignore the multi-level nature of the data as this will produce misleading parameter estimates. There are two levels of interest: school and HEI.

97. Students are nested within HEIs but they are also nested within post-16 schools. It is not true that students who attend the same HEI would necessarily have to have attended the same post-16 school (although it is possible). This means we have multiple hierarchical structures which are not nested within each other. This type of structure is commonly called a cross-classification.

98. To fit this cross-classification structure two additional variance terms are fitted in our original model (given in Equation 1). The first of these terms attempts to identify variation noted in HE achievement associated with the nested-school structure of the data (shown in

Equation 4 as $u^{(2)}_{1,\ school\ id(i)}$). The second of the terms partitions the variation associated with the nested-student-HEI structures ($u^{(3)}_{1,\ hei\ id(i)}$).

99. The cross-classification model is given in Equation 4 (the definitions of Equation 1 also apply to this model):

$$
\left.\begin{array}{l}
\text{firstup}_i \sim \text{Binomial}(\text{denom}_i,\ \pi_i) \\
\text{firstup}_i = \pi_i + e_{0i}\text{bcons}^*_i
\end{array}\right\}
$$

$$
\begin{aligned}
\text{logit}(\pi_i) = {} & \beta_{1i}\text{cons}_i + \beta_2\text{student q}_i + \beta_3\text{school q}_i + \beta_4\text{hei q}_i + \beta_5\text{male}_i + \\
& \beta_6\text{sub 1}_i + \beta_7\text{sub 2}_i + \beta_8\text{sub 3}_i + \beta_9\text{sub 4}_i + \beta_{10}\text{sub 5}_i + \beta_{11}\text{sub 6}_i + \\
& \beta_{12}\text{sub 7}_i + \beta_{13}\text{sub 8}_i + \beta_{14}\text{sub 9}_i + \beta_{15}\text{sub 10}_i + \beta_{16}\text{all girls}_i + \\
& \beta_{17}\text{three year}_i + \beta_{18}\text{lea}_i + \beta_{19}\text{fec}_i + \beta_{20}\text{gms}_i + \beta_{21}\text{selective}_i + \\
& \beta_{22}\text{student q . school q}_i + \beta_{23}\text{student q . male}_i + \\
& \beta_{24}\text{student q . school q . male}_i + \beta_{25}\text{student q . sub 1}_i + \\
& \beta_{26}\text{student q . sub 4}_i + \beta_{27}\text{student q . sub 5}_i + \beta_{28}\text{student q . sub 7}_i + \\
& \beta_{29}\text{student q . sub 8}_i + \beta_{30}\text{student q . male . sub 5}_i + \\
& \beta_{31}\text{student q . gms}_i + \beta_{32}\text{school q . male}_i + \beta_{33}\text{school q . sub 1}_i + \\
& \beta_{34}\text{male . sub 2}_i + \beta_{35}\text{male . sub 4}_i + \beta_{36}\text{male . sub 5}_i + \\
& \beta_{37}\text{male . sub 8}_i + \beta_{38}\text{male . sub 9}_i + \beta_{39}\text{all girls . lea}_i + \\
& \beta_{40}\text{sub 2 . three year}_i + \beta_{41}\text{sub 8 . three year}_i + \beta_{42}\text{sub 9 . three year}_i + \\
& \beta_{43}\text{hei q . three year}_i
\end{aligned}
$$

$$\beta_{1i} = \beta_1 + u^{(3)}_{1,hei\ id(i)} + u^{(2)}_{1,school\ id(i)}$$

$$\left[u^{(3)}_{1,hei\ id(i)}\right] \sim \text{N}(0,\ \Omega^{(3)}_u)\ :\ \Omega^{(3)}_u = \left[\Omega^{(3)}_{u1,1}\right]$$

$$\left[u^{(2)}_{1,school\ id(i)}\right] \sim \text{N}(0,\ \Omega^{(2)}_u)\ :\ \Omega^{(2)}_u = \left[\Omega^{(2)}_{u1,1}\right]$$

$$\text{bcons}^*_i = \text{bcons}_i[\pi_i(1-\pi_i)/\text{denom}_i]^{0.5}$$

$$\left[e_{0i}\right] \sim (0,\ \Omega_e)\ :\ \Omega_e = \left[1\right]$$

(Equation 4)

100. Correctly modelling the nested structures using the cross-classification model does not change the conclusions drawn from the single level model. The cross-classification parameter estimates are given in Table R17 alongside the original single level model estimates.

101. The effect of school performance is estimated to be -0.11 using the single level modelling. The estimated effect is slightly reduced using the cross-classification model, dropping to around -0.08. The effect of each of the school types relative to a independent school student are consistent, with the LEA effect being 0.34 in the single level model and 0.35 in the cross-classification model. The other two school type effects change (single level to cross-classification) from 0.20 to 0.21 for the FEC effect and 0.38 to 0.36 for the grant maintained school type.

<u>Random coefficient models</u>

102. The models described up to this point focus on examining school effects for the population as a whole. It has been concluded that the effect of school performance is inconsistent and small in nature when considered across the whole HEI population. It has also been found that there is a positive effect on HE achievement for LEA schooled students (compared to independent schooling) when the population is considered as a whole. It may well be the case that the effect of both school performance and school type may vary dramatically from HEI to HEI.

103. Random coefficient models allow different parameter estimates between HEIs, which means that we can allow the effect of school performance to vary depending on which HEI the student has attended. A similar approach can be taken for the effect of attending an LEA school compared against an independent school. Other similar models could be considered but are not discussed here.

Random school performance effects for HEIs

104. The first random coefficient model considers allowing the effect of school performance to vary depending on the HEI attended. In this model, the parameter relating to school performance (β_3) has additional HEI random effects terms added to an overall effect, and the remaining structure is as in the cross-classification model. The model is given in Equation 5:

$$
\left.\begin{aligned}
&\text{firstup}_i \sim \text{Binomial}(\text{denom}_i,\ \pi_i) \\
&\text{firstup}_i = \pi_i + e_{0i}\text{bcons}_i^*
\end{aligned}\right\}
$$

$$
\begin{aligned}
\text{logit}(\pi_i) = {} & \beta_{1i}\text{cons}_i + \beta_2\text{student q}_i + \beta_{3i}\text{school q}_i + \beta_4\text{hei q}_i + \beta_5\text{male}_i + \beta_6\text{sub 1}_i + \\
& \beta_7\text{sub 2}_i + \beta_8\text{sub 3}_i + \beta_9\text{sub 4}_i + \beta_{10}\text{sub 5}_i + \beta_{11}\text{sub 6}_i + \beta_{12}\text{sub 7}_i + \\
& \beta_{13}\text{sub 8}_i + \beta_{14}\text{sub 9}_i + \beta_{15}\text{sub 10}_i + \beta_{16}\text{all girls}_i + \beta_{17}\text{three year}_i + \\
& \beta_{18}\text{lea}_i + \beta_{19}\text{fec}_i + \beta_{20}\text{gms}_i + \beta_{21}\text{selective}_i + \beta_{22}\text{student q . school q}_i + \\
& \beta_{23}\text{student q . male}_i + \beta_{24}\text{student q . school q . male}_i + \\
& \beta_{25}\text{student q . sub 1}_i + \beta_{26}\text{student q . sub 4}_i + \beta_{27}\text{student q . sub 5}_i + \\
& \beta_{28}\text{student q . sub 7}_i + \beta_{29}\text{student q . sub 8}_i + \beta_{30}\text{student q . male . sub 5}_i + \\
& \beta_{31}\text{student q . gms}_i + \beta_{32}\text{school q . male}_i + \beta_{33}\text{school q . sub 1}_i + \\
& \beta_{34}\text{male . sub 2}_i + \beta_{35}\text{male . sub 4}_i + \beta_{36}\text{male . sub 5}_i + \beta_{37}\text{male . sub 8}_i + \\
& \beta_{38}\text{male . sub 9}_i + \beta_{39}\text{all girls . lea}_i + \beta_{40}\text{sub 2 . three year}_i + \\
& \beta_{41}\text{sub 8 . three year}_i + \beta_{42}\text{sub 9 . three year}_i + \beta_{43}\text{hei q . three year}_i
\end{aligned}
$$

$$
\beta_{1i} = \beta_1 + u^{(3)}_{1,hei\ id(i)} + u^{(2)}_{1,school\ id(i)}
$$

$$
\beta_{3i} = \beta_3 + u^{(3)}_{3,hei\ id(i)}
$$

$$
\begin{bmatrix} u^{(3)}_{1,hei\ id(i)} \\ u^{(3)}_{3,hei\ id(i)} \end{bmatrix} \sim \text{N}(0,\ \Omega^{(3)}_u)\ :\ \Omega^{(3)}_u = \begin{bmatrix} \Omega^{(3)}_{u1,1} & \\ \Omega^{(3)}_{u3,1} & \Omega^{(3)}_{u3,3} \end{bmatrix}
$$

$$
\begin{bmatrix} u^{(2)}_{1,school\ id(i)} \end{bmatrix} \sim \text{N}(0,\ \Omega^{(2)}_u)\ :\ \Omega^{(2)}_u = \begin{bmatrix} \Omega^{(2)}_{u1,1} \end{bmatrix}
$$

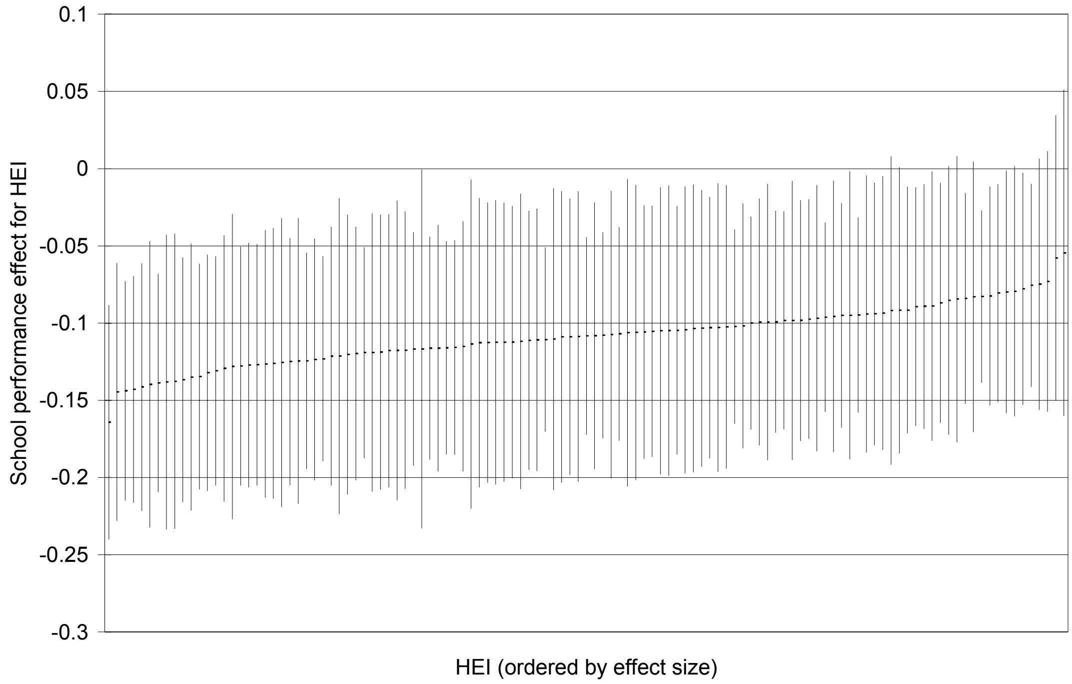
$$
\text{bcons}_i^* = \text{bcons}_i[\pi_i(1-\pi_i)/\text{denom}_i]^{0.5}
$$

$$
\begin{bmatrix} e_{0i} \end{bmatrix} \sim (0,\ \Omega_e)\ :\ \Omega_e = \begin{bmatrix} 1 \end{bmatrix}
$$

(Equation 5)

105. In the cross-classification model, the school performance effect was estimated to be -0.08. In this random coefficient model, the average population effect is now -0.11 but each HEI is allowed to vary from this. The range of HEI effects are show in Figure C2. Each line in the figure represents an HEI with the point in the middle of the line being the estimated HEI effect. The lines show 2 standard deviations above and below the estimated HEI effect.

Figure C2 Variation of the school performance effect for HEIs



106. The HEI variation around the mean school performance effect (0.11) is as expected. Around half the HEIs have smaller school performance effects than the overall average. No HEI displays a behaviour that is significantly different from the standard behaviour.

Random LEA schooling effect for HEIs

107. In this random coefficient model, we allow the effect of attending an LEA school to vary by HEI rather than the effect of school performance. The new model structure is given in Equation 6.

$$\text{firstup}_i \sim \text{Binomial}(\text{denom}_i, \pi_i)$$
$$\text{firstup}_i = \pi_i + e_{0i}\text{bcons}_i^*$$
$$\begin{aligned}\text{logit}(\pi_i) = {} & \beta_{1i}\text{cons}_i + \beta_2\text{student q}_i + \beta_3\text{school q}_i + \beta_4\text{hei q}_i + \beta_5\text{male}_i + \beta_6\text{sub 1}_i + \\ & \beta_7\text{sub 2}_i + \beta_8\text{sub 3}_i + \beta_9\text{sub 4}_i + \beta_{10}\text{sub 5}_i + \beta_{11}\text{sub 6}_i + \beta_{12}\text{sub 7}_i + \\ & \beta_{13}\text{sub 8}_i + \beta_{14}\text{sub 9}_i + \beta_{15}\text{sub 10}_i + \beta_{16}\text{all girls}_i + \beta_{17}\text{three year}_i + \\ & \beta_{18i}\text{lea}_i + \beta_{19}\text{fec}_i + \beta_{20}\text{gms}_i + \beta_{21}\text{selective}_i + \beta_{22}\text{student q . school q}_i + \\ & \beta_{23}\text{student q . male}_i + \beta_{24}\text{student q . school q . male}_i + \\ & \beta_{25}\text{student q . sub 1}_i + \beta_{26}\text{student q . sub 4}_i + \beta_{27}\text{student q . sub 5}_i + \\ & \beta_{28}\text{student q . sub 7}_i + \beta_{29}\text{student q . sub 8}_i + \beta_{30}\text{student q . male . sub 5}_i + \\ & \beta_{31}\text{student q . gms}_i + \beta_{32}\text{school q . male}_i + \beta_{33}\text{school q . sub 1}_i + \\ & \beta_{34}\text{male . sub 2}_i + \beta_{35}\text{male . sub 4}_i + \beta_{36}\text{male . sub 5}_i + \beta_{37}\text{male . sub 8}_i + \\ & \beta_{38}\text{male . sub 9}_i + \beta_{39}\text{all girls . lea}_i + \beta_{40}\text{sub 2 . three year}_i + \\ & \beta_{41}\text{sub 8 . three year}_i + \beta_{42}\text{sub 9 . three year}_i + \beta_{43}\text{hei q . three year}_i\end{aligned}$$
$$\beta_{1i} = \beta_1 + u^{(3)}_{1,hei\ id(i)} + u^{(2)}_{1,school\ id(i)}$$
$$\beta_{18i} = \beta_{18} + u^{(3)}_{18,hei\ id(i)}$$

$$\begin{bmatrix} u^{(3)}_{1,hei\ id(i)} \\ u^{(3)}_{18,hei\ id(i)} \end{bmatrix} \sim \text{N}(0,\ \Omega^{(3)}_u)\ :\ \Omega^{(3)}_u = \begin{bmatrix} \Omega^{(3)}_{u\,1,1} & \\ \Omega^{(3)}_{u\,18,1} & \Omega^{(3)}_{u\,18,18} \end{bmatrix}$$

$$\begin{bmatrix} u^{(2)}_{1,school\ id(i)} \end{bmatrix} \sim \text{N}(0,\ \Omega^{(2)}_u)\ :\ \Omega^{(2)}_u = \begin{bmatrix} \Omega^{(2)}_{u\,1,1} \end{bmatrix}$$

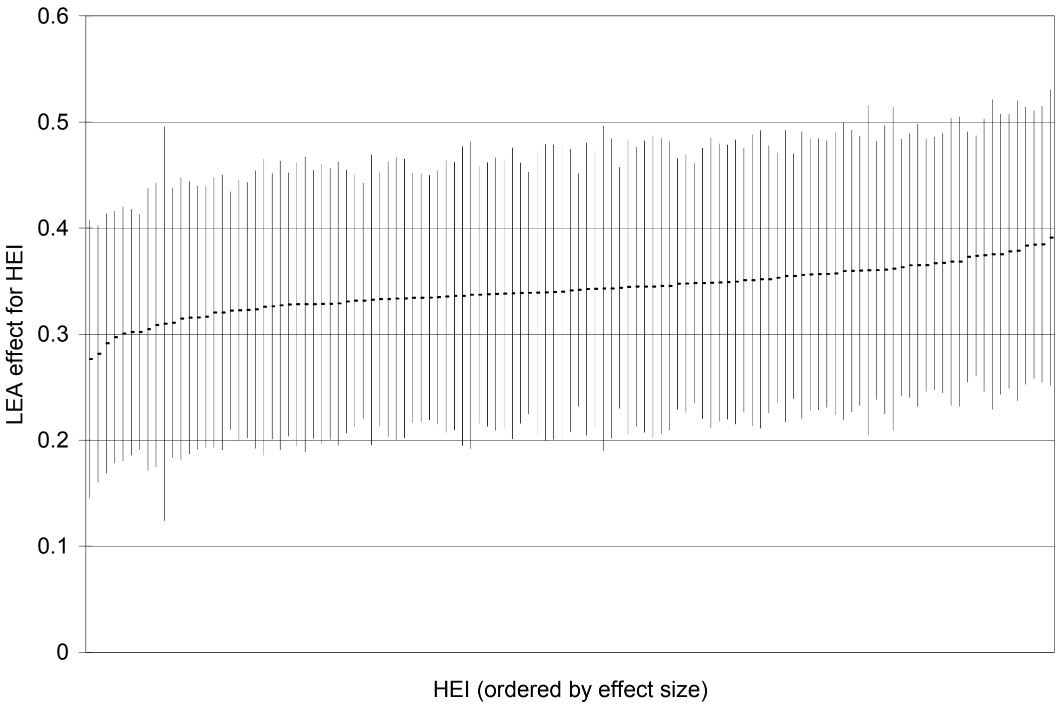
$$\text{bcons}_i^* = \text{bcons}_i[\pi_i(1 - \pi_i)/\text{denom}_i]^{0.5}$$

$$\begin{bmatrix} e_{0i} \end{bmatrix} \sim (0,\ \Omega_e)\ :\ \Omega_e = \begin{bmatrix} 1 \end{bmatrix}$$

(Equation 6)

108. In the original cross-classification model, the LEA effect was estimated to be 0.35. In this random coefficient model, the overall LEA effect is unchanged (0.34). Figure C3 shows a similar figure to Figure C2 but with the HEI LEA effect given instead of the HEI school performance effects. Once again there are no HEIs which show a significantly different LEA effect to the average.

Figure C3 Variation of the LEA effect for HEIs



Random LEA, school performance and student A-level point effects for HEIs

109. The final model that has been fitted allows a random coefficient for the LEA, school performance and student A-level points at the HEI level. This triple random-coefficient cross-classification model is given in Equation 7.

$$
\left.\begin{array}{l}
\text{firstup}_{ijk} \sim \text{Binomial}(\text{denom}_{ijk}, \pi_{ijk}) \\
\text{firstup}_{ijk} = \pi_{ijk} + e_{0ijk}\text{bcons}^{*}
\end{array}\right\}
$$

$$
\begin{aligned}
\text{logit}(\pi_{ijk}) = {} & \beta_{1jk}\text{cons} + \beta_{2k}\text{student q}_{ijk} + \beta_{3k}\text{school q}_{ijk} + \beta_{4}\text{hei q}_{k} + \beta_{5}\text{male}_{ijk} + \\
& \beta_{6}\text{sub 1}_{ijk} + \beta_{7}\text{sub 2}_{ijk} + \beta_{8}\text{sub 3}_{ijk} + \beta_{9}\text{sub 4}_{ijk} + \beta_{10}\text{sub 5}_{ijk} + \\
& \beta_{11}\text{sub 6}_{ijk} + \beta_{12}\text{sub 7}_{ijk} + \beta_{13}\text{sub 8}_{ijk} + \beta_{14}\text{sub 9}_{ijk} + \beta_{15}\text{sub 10}_{ijk} + \\
& \beta_{16}\text{all girls}_{ijk} + \beta_{17}\text{three year}_{ijk} + \beta_{18k}\text{lea}_{ijk} + \beta_{19}\text{fec}_{ijk} + \beta_{20}\text{gms}_{ijk} + \\
& \beta_{21}\text{student q . school q}_{ijk} + \beta_{22}\text{student q . male}_{ijk} + \\
& \beta_{23}\text{student q . school q . male}_{ijk} + \beta_{24}\text{student q . sub 1}_{ijk} + \\
& \beta_{25}\text{student q . sub 4}_{ijk} + \beta_{26}\text{student q . sub 5}_{ijk} + \beta_{27}\text{student q . sub 7}_{ijk} + \\
& \beta_{28}\text{student q . sub 8}_{ijk} + \beta_{29}\text{student q . male . sub 5}_{ijk} + \\
& \beta_{30}\text{student q . gms}_{ijk} + \beta_{31}\text{school q . male}_{ijk} + \beta_{32}\text{school q . sub 1}_{ijk} + \\
& \beta_{33}\text{male . sub 2}_{ijk} + \beta_{34}\text{male . sub 4}_{ijk} + \beta_{35}\text{male . sub 5}_{ijk} + \\
& \beta_{36}\text{male . sub 8}_{ijk} + \beta_{37}\text{male . sub 9}_{ijk} + \beta_{38}\text{all girls . lea}_{ijk} + \\
& \beta_{39}\text{sub 2 . three year}_{ijk} + \beta_{40}\text{sub 8 . three year}_{ijk} + \\
& \beta_{41}\text{sub 9 . three year}_{ijk} + \beta_{42}\text{hei q . three year}_{ijk} + \beta_{43}\text{selective}_{ijk}
\end{aligned}
$$

$$
\begin{aligned}
\beta_{1jk} &= \beta_1 + v_{1k} + u_{1jk} \\
\beta_{2k} &= \beta_2 + v_{2k} \\
\beta_{3k} &= \beta_3 + v_{3k} \\
\beta_{18k} &= \beta_{18} + v_{18k}
\end{aligned}
$$

$$
\begin{bmatrix} v_{1k} \\ v_{2k} \\ v_{3k} \\ v_{18k} \end{bmatrix} \sim \text{N}(0, \Omega_v) : \Omega_v = \begin{bmatrix} \sigma^2_{v1} & & & \\ \sigma_{v21} & \sigma^2_{v2} & & \\ \sigma_{v31} & \sigma_{v32} & \sigma^2_{v3} & \\ \sigma_{v181} & \sigma_{v182} & \sigma_{v183} & \sigma^2_{v18} \end{bmatrix}
$$

$$
\begin{bmatrix} u_{1jk} \end{bmatrix} \sim \text{N}(0, \Omega_u) : \Omega_u = \begin{bmatrix} \sigma^2_{u1} \end{bmatrix}
$$

$$
\text{bcons}^{*} = \text{bcons}[\pi_{ijk}(1 - \pi_{ijk})/\text{denom}_{ijk}]^{0.5}
$$

$$
\begin{bmatrix} e_{0ijk} \end{bmatrix} \sim (0, \Omega_e) : \Omega_e = \begin{bmatrix} 1 \end{bmatrix}
$$

$$
\textit{Deviance(MCMC)} = 91040.130(79005 \text{ of } 79005 \text{ cases in use})
$$

(Equation 7)

110. The overall effect of attending an LEA school is estimated to be 0.34 using this more complex model, compared to 0.35 from the simple single-level model. The overall school performance is also stable at -0.10. Additionally the effect of a student’s A-level points is not

changed using this improved model. Taking these features into account, it is concluded that the main conclusions of the report are unchanged based on this model.

111. Table R18 shows the individual HEI parameters for the three effects that are allowed to vary by HEI. The LEA effect is in the same direction as the overall effect for all the HEIs and remains significant. The school performance effect is occasionally of an opposite sign to the overall school performance effect and is non-significant for a proportion of the HEIs. Additionally, the student A-level effect is always positive for all HEIs.

Subject specific models

Introduction

112. Our focus to this point has been to establish schooling effects at a HE wide level with subject area variation accounted for through interaction terms. We have not attempted to formally examine the variation in schooling effects at a subject area level. It will be true that the schooling effects from subject to subject will vary in strength compared to the overall patterns. In this section, we use a general model specification to examine data at a subject level.

113. This particular analysis will explore the varying strength of the schooling effect within subjects. Therefore for simplicity's sake we concentrate on the five main explanatory variables: student entry qualifications; HEI quality; gender; school type; and school performance. For illustration, we analyse students who studied mathematics in the section below.

Mathematics

114. In the data there were 6,632 students recorded as studying mathematics in 1997-98, coming from 1,733 schools and attending 86 different HEIs. The single level model given in Equation 8 was fitted to these mathematics data.

$$
\left.\begin{array}{l}
\text{firstup}_{ijk} \sim \text{Binomial}(\text{denom}_{ijk}, \pi_{ijk}) \\
\text{firstup}_{ijk} = \pi_{ijk} + e_{0ijk}\text{bcons}^{*}
\end{array}\right\}
$$

$$
\begin{aligned}
\text{logit}(\pi_{ijk}) = {} & \beta_1\text{cons} + \beta_2\text{student q}_{ijk} + \beta_3\text{school q}_{ijk} + \beta_4\text{hei q}_k + \beta_5\text{male}_{ijk} + \\
& \beta_6\text{lea}_{ijk} + \beta_7\text{fec}_{ijk} + \beta_8\text{gms}_{ijk} + \beta_9\text{student q . hei q}_{ijk} + \\
& \beta_{10}\text{student q . gms}_{ijk} + \beta_{11}\text{school q . hei q}_{ijk} + \beta_{12}\text{hei q . male}_{ijk} + \\
& \beta_{13}\text{male . fec}_{ijk}
\end{aligned}
$$

$$
\text{bcons}^{*} = \text{bcons}[\pi_{ijk}(1 - \pi_{ijk})/\text{denom}_{ijk}]^{0.5}
$$

$$
\left[e_{0ijk}\right] \sim (0,\ \Omega_e) : \Omega_e = \left[1\right]
$$

(Equation 8)

115. This single level model ignores the cross-classification structure of the data and does not allow key effects to vary by HEI. We can improve and develop this model to allow for these features using the models described previously.

116. Potentially there are many models that could be examined including additional effects and additional structures. We concentrate on two of these (Equations 9 and 10). Both take into account the multi-level cross-classification nature of the mathematics data allowing for the assumption that students can attend different HEIs and different post-16 schools. The difference between the two models is how the HEI variation in the explanatory terms is dealt with. In Equation 9 it is assumed that any HEI variation in the key explanatory terms is related to HEI quality (as is assumed in the single level mathematics model).

$$
\left.\begin{aligned}
&\text{firstup}_i \sim \text{Binomial}(\text{denom}_i,\ \pi_i)\\
&\text{firstup}_i = \pi_i + e_{0i}\text{bcons}_i^*
\end{aligned}\right\}
$$

$$
\begin{aligned}
\text{logit}(\pi_i) = {} & \beta_{1i}\text{cons}_i + \beta_2\text{student q}_i + \beta_3\text{school q}_i + \beta_4\text{hei q}_i + \beta_5\text{male}_i + \\
& \beta_6\text{lea}_i + \beta_7\text{fec}_i + \beta_8\text{gms}_i + \beta_9\text{student q . hei q}_i + \\
& \beta_{10}\text{student q . gms}_i + \beta_{11}\text{school q . hei q}_i + \beta_{12}\text{hei q . male}_i + \\
& \beta_{13}\text{male . fec}_i
\end{aligned}
$$

$$
\beta_{1i} = \beta_1 + u^{(3)}_{1,hei\ id(i)} + u^{(2)}_{1,school\ id(i)}
$$

$$
\left[u^{(3)}_{1,hei\ id(i)}\right] \sim \text{N}(0,\ \Omega^{(3)}_u)\ :\ \Omega^{(3)}_u = \left[\Omega^{(3)}_{u1,1}\right]
$$

$$
\left[u^{(2)}_{1,school\ id(i)}\right] \sim \text{N}(0,\ \Omega^{(2)}_u)\ :\ \Omega^{(2)}_u = \left[\Omega^{(2)}_{u1,1}\right]
$$

$$
\text{bcons}_i^* = \text{bcons}_i[\pi_i(1 - \pi_i)/\text{denom}_i]^{0.5}
$$

$$
\left[e_{0i}\right] \sim (0,\ \Omega_e)\ :\ \Omega_e = \left[1\right]
$$

$$
Deviance(MCMC) = 8013.772(6632 \text{ of } 6632 \text{ cases in use})
$$

(Equation 9)

117. The alternative (Equation 10) is to allow HEI quality as a main effect but any variation in the other explanatory effects is dealt with through random coefficients for the key main effects.

$$
\left.\begin{array}{l}
\text{firstup}_i \sim \text{Binomial}(\text{denom}_i, \pi_i) \\
\text{firstup}_i = \pi_i + e_{0i}\text{bcons}_i^*
\end{array}\right\}
$$

$$
\begin{aligned}
\text{logit}(\pi_i) = {} & \beta_{1i}\text{cons}_i + \beta_{2i}\text{student q}_i + \beta_{3i}\text{school q}_i + \beta_4\text{hei q}_i + \beta_{5i}\text{male}_i + \\
& \beta_6\text{lea}_i + \beta_7\text{fec}_i + \beta_8\text{gms}_i + \beta_9\text{student q . gms}_i + \beta_{10}\text{male . fec}_i
\end{aligned}
$$

$$
\begin{aligned}
\beta_{1i} &= \beta_1 + u^{(3)}_{1,hei\ id(i)} + u^{(2)}_{1,school\ id(i)} \\
\beta_{2i} &= \beta_2 + u^{(3)}_{2,hei\ id(i)} \\
\beta_{3i} &= \beta_3 + u^{(3)}_{3,hei\ id(i)} \\
\beta_{5i} &= \beta_5 + u^{(3)}_{5,hei\ id(i)}
\end{aligned}
$$

$$
\begin{bmatrix} u^{(3)}_{1,hei\ id(i)} \\ u^{(3)}_{2,hei\ id(i)} \\ u^{(3)}_{3,hei\ id(i)} \\ u^{(3)}_{5,hei\ id(i)} \end{bmatrix} \sim \text{N}(0,\ \Omega^{(3)}_u)\ :\ \Omega^{(3)}_u = \begin{bmatrix} \Omega^{(3)}_{u1,1} & & & \\ \Omega^{(3)}_{u2,1} & \Omega^{(3)}_{u2,2} & & \\ \Omega^{(3)}_{u3,1} & \Omega^{(3)}_{u3,2} & \Omega^{(3)}_{u3,3} & \\ \Omega^{(3)}_{u5,1} & \Omega^{(3)}_{u5,2} & \Omega^{(3)}_{u5,3} & \Omega^{(3)}_{u5,5} \end{bmatrix}
$$

$$
\left[u^{(2)}_{1,school\ id(i)}\right] \sim \text{N}(0,\ \Omega^{(2)}_u)\ :\ \Omega^{(2)}_u = \left[\Omega^{(2)}_{u1,1}\right]
$$

$$
\text{bcons}_i^* = \text{bcons}_i[\pi_i(1 - \pi_i)/\text{denom}_i]^{0.5}
$$

$$
\left[e_{0i}\right] \sim (0,\ \Omega_e)\ :\ \Omega_e = \left[1\right]
$$

$$
Deviance(MCMC) = 7963.566(6632 \text{ of } 6632 \text{ cases in use})
$$

(Equation 10)

118. The coefficients of these models are given in Table R19.

119. We can compare the mathematics parameter estimates for the school performance and school type to the equivalent parameter estimates if the model in Equation 10 is run on the whole dataset. This approach gives us a rough idea of how the schooling effects vary within mathematics compared to the population as a whole. This model is favoured because the parameter effects are easier to interpret due to the lack of interaction terms.

120. One approach is to look at what the average effect is of school performance and of changing an LEA student to an independent school student. The parameter estimate of school performance for the overall data is -0.04 for every school A-level point. If a student's school performance was dropped by two points (equivalent to the analysis in Table 3 of the main report), the logit parameter would be increased by 0.07. An increase of 0.5 A-level

points would be required to cause the same change in the logit parameter. This gives us a rough measure of how much effect school performance has in the overall data.

121. 0.34 is the parameter change associated with changing a student's school type from independent to LEA in the overall data. A change of 2.4 in the student's A-level points would cause the same change in the logit parameter.

122. The LEA effect is roughly equivalent to 9.5 A-level points (or around five A-level grades, that is, the whole range of school performances) decrease in school performance.

123. An equivalent analysis of the mathematics model shows that the school performance effect is larger in mathematics than in the general population. The comparable student A-level effect is 1.0. The effect of changing a student from an independent school student to an LEA student is also weaker in mathematics in comparison to the effect for the whole cohort (1.6 compared to 3.1). The LEA effect for mathematics is equivalent to a drop of 3.1 A-level points in school performance (a weaker effect of school type than seen in the whole cohort's data).

Model results for each subject area

124. Similar analyses have been carried out for the other subject areas. The model parameters for these different analyses are given in Table R20. The implied school performance and LEA effects in terms of implied student A-level points are given in Table C9.

125. In most of the subject areas, the same model structure as in Equation 10 has been fitted but in a small number of cases, a slightly simplified model is fitted due to a lack of data and/or effect of certain parameters. For example, the model for social sciences does not fit a term for FECs as it is near zero.

Table C9 Variation in schooling effects by subject area

Subject area	N	School performance	LEA school	Relative LEA strength
Allied to medicine	3,499	1.6	2.4	-2.9
Biological or physical sciences	14,448	0.4	2.7	-13.1
Agriculture	606	-0.5	2.5	9.4
Mathematical sciences	6,632	1.0	1.6	-3.1
Engineering	4,621	1.6	2.2	-2.9
Social studies	11,311	0.5	2.7	-11.8
Business	8,985	1.0	3.8	-7.6
Languages	11,955	-0.4	1.5	7.6
Creative arts	2,906	1.4	1.2	-1.8
Education	3,757	1.2	3.0	-4.9
Combined studies	10,285	0.4	1.4	-7.5
Overall	79,005	0.5	2.4	-9.5

126. Table C9 shows that there is some variation in the strength of the school performance and LEA school effects between subject areas. In the largest subject areas, the relative strength of the LEA school effect compared to the school performance effect is around the same as in the overall data.

127. For two of the subject areas (languages and agriculture, which have very low student numbers) the school performance effect is negative, meaning that attending a higher quality school improved the student's HE achievement.

128. The weakest LEA school effects relative to school performance are seen in: the creative arts, where attending an LEA school gives the student only a 1.2 A-level point advantage over their independent school counterparts; and in engineering and subjects allied to medicine, where school performance effects are stronger (1.6 A-level points compared to the overall average of 0.5 A-level points).

129. The conclusion of the subject area analysis is that the LEA school effect is present in all the subject areas. But the school performance effects are inconsistent, being small in general but stronger in certain subjects.

References

Collett, D (1991), *Modelling Binary Data*, London: Chapman and Hall

Cox, D R and Snell, E J (1989), *The Analysis of Binary Data*, Second Edition, London: Chapman and Hall

Goldstein, H (1995), *Multi-level statistical models*, Second Edition, London: Edward Arnold

Rasbash, J, Browne, W, Goldstein, H, Yang, M, et al. (2000). *A user's guide to MLwiN*, Second Edition, London: Institute of Education

SAS (1999). SAS system for Windows (Release 8.0) SAS Inst., Inc. Cary, NC